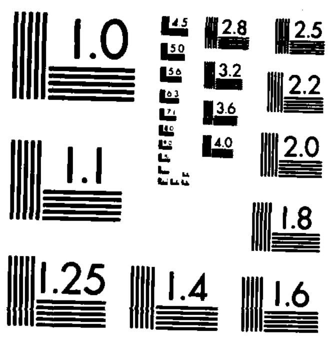
AD-A155 056 DECOMPOSING THE JOINT DISTRIBUTION OF SEVERAL LINEAR COMBINATIONS OF UNIFORM SPACINGS(U) STANFORD UNIV CA DEPT OF STATISTICS F HUFFER 25 APR 85 TR-356 1/1

UNCLASSIFIED N00014-76-C-0475 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A155 055

(2)

DECOMPOSING THE JOINT DISTRIBUTION OF SEVERAL LINEAR COMBINATIONS OF UNIFORM SPACINGS

BY

FRED HUFFER

TECHNICAL REPORT NO. 356
APRIL 25, 1985

PREPARED UNDER CONTRACT
N00014-76-C-0475 (NR-042-267)
FOR THE OFFICE OF NAVAL RESEARCH

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

DTIC
ELECTE
JUN 17 1985
S D
B

DTIC FILE COPY

85 5 24 201

DECOMPOSING THE JOINT DISTRIBUTION OF SEVERAL LINEAR COMBINATIONS OF UNIFORM SPACINGS

BY

FRED HUFFER

TECHNICAL REPORT NO. 356

APRIL 25, 1985

Prepared Under Contract

N00014-76-C-0475 (NR-042-267)

For the Office of Naval Research

Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government

Approved for public release; distribution unlimited.

DTIC
S ELECTE D
JUN 17 1985
B

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

DECOMPOSING THE JOINT DISTRIBUTION OF SEVERAL

LINEAR COMBINATIONS OF UNIFORM SPACINGS

By

Fred Huffer

1. The Main Result.

Place $n-1$ independent uniformly distributed points on the unit interval. These points divide the unit interval into n pieces whose lengths (called spacings) are denoted $S_1, S_2, \ldots, S_n$. The random vector $\underset{\sim}{S} = (S_1, S_2, \ldots, S_n)'$ is uniformly distributed on the simplex determined by $S_1 + S_2 + \cdots + S_n = 1$. It follows that the S_i are exchangeable random variables. For further information on spacings see Pyke (1965).

Let Γ be a $k \times n$ real matrix. We wish to investigate the distribution of the random vector $\Gamma\underset{\sim}{S}$. Let $\mathbb{P}(\Gamma)$ denote the probability measure of $\Gamma\underset{\sim}{S}$ so that $\mathrm{Prob}\{\Gamma\underset{\sim}{S} \varepsilon A\} = (\mathbb{P}(\Gamma))(A)$. We shall present an identity which in some cases allows us to express $\mathbb{P}(\Gamma)$ as a sum of simpler terms.

For any $\xi \varepsilon \mathbb{R}^k$, define $\Gamma_{i,\xi}$ to be the $k \times n$ matrix obtained by replacing the $i^{\underline{th}}$ column of Γ by ξ. Let $e \varepsilon \mathbb{R}^n$ be the vector of ones, $e' = (1,1,\ldots,1)$. Our main result is the following.

Theorem: Suppose $d' = (d_1, d_2, \ldots, d_n)$ satisfies $e'd = 1$. Let $\xi = \Gamma d$. Then

$$\mathbb{P}(\Gamma) = \sum_{i=1}^{n} d_i \mathbb{P}(\Gamma_{i,\xi}).$$

This theorem is proved in section 2. Section 3 contains many examples using the theorem to find probabilities in situations involving one or several linear combinations of spacings. Other approaches to these problems have been given in the literature. A number of authors have obtained the distribution of a single linear combination of spacings which we deal with in example 1. Finding this distribution is equivalent to determining the volume of the intersection of a simplex with a half-space. Dempster and Kleyle (1968) employ an inclusion-exclusion argument to decompose this region into simplices and thus obtain the volume in the case when the coefficients of the linear combination are all different. Weisberg (1971) generalized their result to arbitrarily coincident coefficients. Varsi (1973) gave an algorithm for computing this volume based on a different method of dissecting the region into simplices. Ali (1973) obtains the distribution by inversion of its characteristic function. A more elementary analytic argument (avoiding inversion of transforms) is given by Gerber (1981). Other closely related results concern the distribution of a linear combination of order statistics from the uniform distribution (Ali (1969)) and the distribution of a linear combination of independent exponential random variables (Ali and Obaidullah (1982), Mathai (1983)). The papers by Ali (1973), Gerber (1981) and Ali and Obaidullah (1982) all make use of properties of divided differences as does the present paper.

The problem of finding the joint distribution of several linear combinations of uniform spacings has received less attention. Some special cases were dealt with by Kleyle (1967); see the remarks in section 3 of Dempster and Kleyle (1968) and Theorem A.2 of Kleyle (1971). The closely

related problem of finding the joint distribution of linear combinations of order statistics from the uniform distribution was considered by Ali and Mead (1969) who obtained results by inverting the joint characteristic function.

There is a large body of work dealing with probabilities of clustering on the line or circle, probabilities of multiple coverage of the line or circle, and the distribution of the scan statistic and other tests for uniformity. Specialized methods have been developed to calculate these probabilities. See Naus (1982) for references. These problems frequently involve calculating $P\{X_i > t \text{ for } 1 \le i \le k\}$ or $P\{X_i < t \text{ for } 1 \le i \le k\}$ where $(X_1, X_2, \ldots, X_k)' = \Gamma \underset{\sim}{S}$. Here Γ is some matrix whose entries are all zeros and ones which are usually arranged according to a simple pattern. For example, Glaz and Naus (1983) find the variance of the number of clusters on the line by calculating

$$(*) \qquad P\left\{ \sum_{i=1}^{k} S_i < t,\ \sum_{i=j}^{j+k-1} S_i < t \right\}$$

for arbitrary j, k and t. Examples 3 through 6 illustrate how our theorem may be used in these situations. In particular, example 3 provides another route to the calculation of (*) besides that given in Theorem 1 of Glaz and Naus (1983).

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

2. Proof of the Theorem.

The proof uses the notation of divided differences which is presented in many texts on numerical analysis or the calculus of finite differences. Two good sources are the texts by Gel'fond and Milne-Thomson.

For any function f define

$$f[x_1,x_2] = \frac{f(x_2)-f(x_1)}{x_2-x_1} .$$

For $n > 2$ the divided difference is defined inductively by

$$f[x_1,x_2,\ldots,x_n] = \frac{f[x_2,x_3,\ldots,x_n]-f[x_1,x_2,\ldots,x_{n-1}]}{x_n-x_1} .$$

We shall need the following facts. When $x_1,x_2,\ldots,x_n$ are all distinct we may write

(1) $$f[x_1,x_2,\ldots,x_n] = \sum_{i=1}^{n} \left(\frac{f(x_i)}{\prod_{j\neq i} (x_i-x_j)} \right) .$$

If f has a continuous $(n-1)$th derivative $f^{(n-1)}$, then for any $x_1,x_2,\ldots,x_n$ we have

(2) $$(n-1)!f[x_1,x_2,\ldots,x_n] = Ef^{(n-1)}\left(\sum_{i=1}^{n} x_i S_i\right) .$$

Formula (2) is an immediate consequence of a result due to Hermite which may be found in section 1.6 of Milne-Thomson.

We shall abbreviate the notation for divided differences. For any vector $x = (x_1, x_2, \ldots, x_n)'$ we let $f[x] = f[x_1, x_2, \ldots, x_n]$. Taking $f(u) = e^u$ (written as exp) in formula (2) gives an expression for the Laplace transform of $\underset{\sim}{S}$ which is

$$\text{(3)} \qquad E \exp(\theta' \underset{\sim}{S}) = (n-1)! \exp[\theta]$$

where $\theta = (\theta_1, \theta_2, \ldots, \theta_n)'$. The Laplace transform of $\Gamma \underset{\sim}{S}$ may be obtained from (3) by inspection and is

$$\text{(4)} \qquad E \exp(\theta'(\Gamma \underset{\sim}{S})) = (n-1)! \exp[\Gamma' \theta]$$

where $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$.

Our theorem follows directly from (4) and the next lemma on divided differences. For any vector $x = (x_1, x_2, \ldots, x_n)'$ and any real number u, define $x_{i,u}$ to be the vector obtained by replacing the $i^{\underline{th}}$ component of x by u.

<u>Basic Lemma</u>: *Let f by any function having n-1 continuous derivatives. Suppose $d = (d_1, d_2, \ldots, d_n)'$ satisfies $e'd = 1$. For all $x \in \mathbb{R}^n$ we have*

$$f[x] = \sum_{i=1}^{n} d_i f[x_{i,u}]$$

where $u = x'd$.

When $u, x_1, x_2, \ldots x_n$ are all distinct, the lemma may be verified by routine algebra after expanding all the divided differences using formula

(1). From (2) it is clear that $f[x_1,x_2,\ldots,x_n]$ is a continuous function of its arguments so that the lemma must hold for all $x_1,x_2,\ldots,x_n$.

Let $\xi = \Gamma d$ and $u = \theta'\xi$. Noting that $(\Gamma'\theta)_{i,u} = (\Gamma_{i,\xi})'\theta$ we conclude from the lemma that

$$\exp[\Gamma'\theta] = \sum_{i=1}^{n} d_i \exp[(\Gamma_{i,\xi})'\theta] .$$

Thus we can express the Laplace transform in (4) as a linear combination of similar Laplace transforms. This proves the theorem.

3. Examples.

Example 1: When Γ is $1 \times n$ we can explicitly calculate the distribution of $\Gamma \underset{\sim}{S}$. Choose $k \leq n$ and let $a_1,a_2,\ldots,a_k$ be any real numbers. We shall apply the theorem with $\Gamma = (a_1,a_2,\ldots,a_n)$ where $a_i = 0$ for $i > k$.

Assume $a_1 \neq a_2$. Define

$$d_1 = \frac{a_2}{a_2-a_1} \quad \text{and} \quad d_2 = \frac{a_1}{a_1-a_2} .$$

Clearly $d_1+d_2 = 1$ and $d_1a_1+d_2a_2 = 0$. Thus $\mathbb{P}(\Gamma) = d_1\mathbb{P}(\Gamma_{1,0})+d_2\mathbb{P}(\Gamma_{2,0})$. For any $t > 0$, this implies that

$$P\{\sum_i a_iS_i > t\} = d_1P\{\sum_{i\neq 1} a_iS_i > t\} + d_2P\{\sum_{i\neq 2} a_iS_i > t\}$$

where the sums are taken over $1 \leq i \leq k$ except as indicated. Fix $t > 0$ and define

$$Q(a_1,a_2,\ldots,a_k) = P\{\sum_i a_i S_i > t\} .$$

The previous formula may now be written as

$$(5) \qquad Q(a_1,a_2,\ldots,a_k) = \frac{a_2 Q(a_2,a_3,\ldots,a_k) - a_1 Q(a_1,a_3,\ldots,a_k)}{a_2 - a_1} .$$

The number of arguments in Q is permitted to vary. Let $(z)_+$ denote the positive part, $(z)_+ = \max(z,0)$. It is easy to show that

$$(6) \qquad Q(a) = P\{aS_1 > t\} = (1 - \frac{t}{a})_+^{n-1} .$$

Thus when $a_1,a_2,\ldots,a_k$ are all distinct, repeated application of (5) leads to an explicit formula for $Q(a_1,a_2,\ldots,a_k)$. This formula can be written as a divided difference. To see this we need the next fact.

(7) Lemma: Suppose $s,t,x_1,x_2,\ldots,x_m$ are distinct real numbers. We use the symbol $\underset{\sim}{x}$ to stand for the list $x_1,x_2,\ldots,x_m$. Let the functions f and g satisfy $g(x) = xf(x)$ for all x. Then

$$g[s,t,\underset{\sim}{x}] = \frac{sf[s,\underset{\sim}{x}] - tf[t,\underset{\sim}{x}]}{s-t} .$$

The proof of (7) follows by elementary algebra after expanding the divided differences using formula (1).

From (5) and (6) we obtain

$$Q(a_1,a_2) = \frac{a_2(1 - \frac{t}{a_2})_+^{n-1} - a_1(1 - \frac{t}{a_1})_+^{n-1}}{a_2 - a_1} = f[a_1,a_2]$$

where $f(a) = a(1 - \frac{t}{a})_+^{n-1}$. Using (5) and (7), it now follows by induction that

(8) $$Q(a_1, a_2, \ldots, a_k) = g[a_1, a_2, \ldots, a_k]$$

with $g(a) = a^{k-1}(1 - \frac{t}{a})_+^{n-1}$. This formula has been demonstrated for $k \leq n$ when the values $a_1, a_2, \ldots, a_k$ are distinct, but it is easily shown to hold for all values $a_1, a_2, \ldots, a_k$ except in the single case when $k = n$ and $a_1 = a_2 = \cdots = a_n = t$. The divided difference is not defined for this case.

Numerical analysis texts contain prescriptions for evaluating $g[a_1, a_2, \ldots, a_k]$ when some of the arguments coincide. For instance, when $a_1 = a_2 = \cdots = a_k = a$ and g has $k-1$ continuous derivatives (denoted $g^{(i)}$), then

(9) $$g[a_1, a_2, \ldots, a_k] = \frac{g^{(k-1)}(a)}{(k-1)!} .$$

This is a standard result which follows immediately from formula (2).

To use these prescriptions we must be able to evaluate the derivatives of g. It is not difficult to verify by induction that

(10) $$\frac{1}{r!}\left(\frac{d}{dx}\right)^r [x^m (b - \frac{t}{x})^p] = x^{m-r}\left[\sum_{i=0}^{r} \binom{m-i}{r-i}\binom{p}{i}\left(\frac{t}{x}\right)^i (b - \frac{t}{x})^{p-i}\right]$$

and in particular obtain

(11) $$\frac{1}{m!}\left(\frac{d}{dx}\right)^m [x^m (b - \frac{t}{x})^p] = \sum_{i=0}^{m} \binom{p}{i}\left(\frac{t}{x}\right)^i (b - \frac{t}{x})^{p-i} .$$

Using (8), (9) and (11) we immediately obtain

$$P\{\sum_{i=1}^{k} S_i > t\} = \sum_{i=0}^{k-1} \binom{n-1}{i} t^i (1-t)^{n-1-i} .$$

This formula can also be obtained by a simple direct argument which, however, does not apply in more complicated situations.

Note: Let $Z_1, Z_2, \ldots, Z_k$ be i.i.d. exponential random variables with density e^{-z}. If we let $n \to \infty$ holding k fixed, it is easy to see that $(nS_1, nS_2, \ldots, nS_k)$ converges in distribution to $(Z_1, Z_2, \ldots, Z_k)$. Therefore, substituting $t = x/n$ in formula (8) and letting $n \to \infty$ yields

$$(12) \qquad P\{\sum_{i=1}^{k} a_i Z_i > x\} = f[a_1, a_2, \ldots, a_k]$$

where $f(a) = a^{k-1} e^{-x/a}$.

Partial Divided Differences.

In order to generalize the result of example 1 we need to introduce divided differences for functions with more than one argument. The definitions will be presented in terms of a function f having two arguments, $f = f(x|y)$. The definitions are similar for functions with more than two arguments. Let $\underset{\sim}{u}$ and $\underset{\sim}{w}$ represent arbitrary lists of real numbers $u_1, u_2, \ldots, u_p$ and $w_1, w_2, \ldots, w_q$. Define the divided differences recursively by

$$f[x_1, x_2|y] = \frac{f(x_2|y) - f(x_1|y)}{x_2 - x_1} ,$$

$$f[x|y_1,y_2] = \frac{f(x|y_2)-f(x|y_1)}{y_2-y_1} ,$$

$$f[x_1,x_2,\underset{\sim}{u}|\underset{\sim}{w}] = \frac{f[x_2,\underset{\sim}{u}|\underset{\sim}{w}]-f[x_1,\underset{\sim}{u}|\underset{\sim}{w}]}{x_2-x_1} ,$$

and

$$f[\underset{\sim}{u}|y_1,y_2,\underset{\sim}{w}] = \frac{f[\underset{\sim}{u}|y_2,\underset{\sim}{w}]-f[\underset{\sim}{u}|y_1,\underset{\sim}{w}]}{y_2-y_1} .$$

It may be shown that the divided differences do not depend on the order in which the differences are taken. In particular, for each fixed value of x we may define the function f_x by $f_x(y) = f(x|y)$ and write

(13) $$f[x_1,x_2,\ldots,x_p|y_1,y_2,\ldots,y_q]$$
$$= g[x_1,x_2,\ldots,x_p] \quad \text{where}$$
$$g(x) = f_x[y_1,y_2,\ldots,y_q] .$$

When the values $x_1,x_2,\ldots,x_p$ are distinct and $y_1,y_2,\ldots,y_q$ are also distinct, the divided difference is given by the formula

(14) $$f[x_1,\ldots,x_p|y_1,\ldots,y_q]$$
$$= \sum_{i=1}^{p} \sum_{j=1}^{q} \left(\frac{f(x_i|y_j)}{\prod\limits_{k\neq i} (x_i-x_k) \prod\limits_{\ell\neq j} (y_j-y_\ell)} \right) .$$

REFERENCES

Ali, M.M. (1969). Distribution of linear combination of order statistics from rectangular population. Bull. Inst. Statist. Res. Tr. 3, 1 - 21.

Ali, M.M. and Mead, E.R. (1969). On the distribution of several linear combinations of order statistics from the uniform distribution. Bull. Inst. Statist. Res. Tr. 3, 22 - 41.

Ali, M.M. (1973). Content of the frustum of a simplex. Pacific J. Math. 48, 313 - 322.

Ali, M.M. and Obaidullah, M. (1982). Distribution of linear combination of exponential variates. Commun. Statist.-Theor. Meth. 11, 1453 - 1463.

Cressie, N. (1977). On some properties of the scan statistic on the circle and the line. Journal of Applied Probability, 14, 272 - 283.

Dempster, A.P. and Kleyle, R.M. (1968). Distributions determined by cutting a simplex with hyperplanes. Ann. Math. Statist. 39, 1473 - 1478.

Feller, W. (1971). An Introduction to Probability Theory and Its Applications, Vol. II (Second edition). Wiley, New York.

Gel'fond, A.O. (1971). Calculus of Finite Differences. Hindustan Publishing Corporation, Delhi-7, India.

Gerber, L. (1981). The volume cut off a simplex by a half-space. Pacific J. Math. 94, 311 - 314.

Glaz, J. and Naus, J. (1983). Multiple clusters on the line. Commun. Statist.- Theor. Meth. 12, 1961 - 1986.

Huffer, F. (1982). The Moments and Distributions of Some Quantities Arising from Random Arcs on the Circle. Ph.D. dissertation, Department of Statistics, Stanford University.

Kleyle, R.M. (1967). An application of the direct probability method of inference. Ph.D. thesis, Department of Statistics, Harvard University.

Kleyle, R.M. (1971). Upper and lower posterior probabilities for truncated means. Ann. Math. Statist. 42, 976 - 990.

Mathai, A.M. (1983). On linear combinations of independent exponential variables. Commun. Statist.- Theor. Meth. 12, 625 - 632.

Acknowledgment

I wish to thank Herbert Solomon both for his support of this work and because this work was begun in section 6 of my thesis under his supervision. I thank Persi Diaconis for suggesting the usefulness of divided differences and bringing to my attention formulas (3) and (12).

$$\begin{Bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{Bmatrix} = \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{Bmatrix} .$$

Putting this all together gives

$$\{A\} = 2 \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{Bmatrix} + 2 \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{Bmatrix} - 2 \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} .$$

Evaluating these terms using formula (22) leads to the final answer

$$\{A\} = (1-2t)_+^{n-1} + 2(n-1)t(1-2t)_+^{n-2} + 4\binom{n-1}{2}t^2(1-2t)_+^{n-3} + 4\binom{n-1}{3}t^3(1-2t)_+^{n-4}$$

$$+ 2\binom{n-1}{4}t^4(1-2t)_+^{n-5}.$$

Tables given by Naus (1966) provide a numerical check for this answer. Naus' notation differs from ours. Using our formula (with $n=8$) to calculate $1-\{A\}_t$ for the values $t = .10, .20, .30, \ldots$ leads to the numbers given in Naus' Table 2 for the case $N = 7$, $n = 4$.

Note: As in the note after example 1, we can transform our results about spacings into results about i.i.d. exponentials by substituting $t = x/n$ and letting $n \to \infty$. Let $Z_1, Z_2, \ldots, Z_6$ be i.i.d. with density e^{-z}. The result of example 6 implies that

$$P\{Z_1+Z_2+Z_3 > x, Z_2+Z_3+Z_4 > x, Z_3+Z_4+Z_5 > x, Z_4+Z_5+Z_6 > x\}$$

$$= e^{-2x}\{1+2x+4(x^2/2!) + 4(x^3/3!) + 2(x^4/4!)\}$$

When $n=8$, $(S_1,S_2,\ldots,S_6)$ can be viewed as the spacings between 7 points placed uniformly and independently on the unit interval. (The 7 points divide the unit interval into $n=8$ pieces, but we are for the moment ignoring the leftmost and rightmost pieces.) $\{A\}_t$ is the probability that no interval of length t contains more than 3 of the 7 points.

The columns of A satisfy $\gamma_1-\gamma_3+\gamma_4 = \gamma_6 = (0,0,0,1)'$ so that our theorem applies to yield

$$\{A\} = \begin{Bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{Bmatrix} .$$

These terms have been simplified and rearranged using (a) through (d). Call the above three matrices B, C and D respectively so that $\{A\} = \{B\} - \{C\} + \{D\}$. The columns of each of the matrices B, C and D satisfy $\gamma_1-\gamma_2+\gamma_3 = (0,0,1)'$. Applying our theorem thus leads to

$$\{B\} = \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{Bmatrix} ,$$

$$\{C\} = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{Bmatrix} \quad \text{and}$$

$$\{D\} = \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{Bmatrix} .$$

All of these terms are simple except the third term in $\{B\}$ which again satisfies $\gamma_1-\gamma_2+\gamma_3 = (0,0,1)'$ so that

evaluated using formulas (20) or (21). However, it is convenient to further simplify some of the terms.

$$\begin{Bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{Bmatrix} = 2 \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{Bmatrix}$$

and

$$\begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{Bmatrix} = 2 \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix}$$

because in both cases $2\gamma_4 - \gamma_3 = (0,0,0)'$. Putting all of these results back in our original expression gives

$$\{A\} = 16 \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{Bmatrix} - 10 \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix} - 5 \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{Bmatrix} .$$

Using formulas (20) or (22) to evaluate these terms, we obtain

$$\{A\} = 16(1-(5/2)t)_+^{n-1} - 15(1-3t)_+^{n-1} - 5(n-1)t(1-3t)_+^{n-2} .$$

<u>Example 6</u>:

Our final example will be to compute $\{A\}_t$ for

$$A = \begin{Bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{Bmatrix} .$$

γ_i will be used generically to denote the $i^{\underline{th}}$ column of a matrix. We begin by noting that

$$1-1+1-1+1 = 1 \quad \text{and}$$

$$\gamma_1-\gamma_2+\gamma_3-\gamma_4+\gamma_5 = (0,0,0,0,2)'$$

so that our theorem expresses $\{A\}$ as an alternating sum of five terms, successively replacing each column by $(0,0,0,0,2)'$. By using properties (a) and (d) to simplify these terms and then combining terms which are seen to be equal upon using properties (b) and (c), we obtain

$$\{A\} = 2\begin{Bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 0 & 1 \end{Bmatrix} - 2\begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{Bmatrix} + \begin{Bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 1 \end{Bmatrix} .$$

Noting that the columns of the first term satisfy $\gamma_3-\gamma_4+\gamma_5 = (0,0,0,1)'$, the first term can be rewritten as

$$\begin{Bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 0 & 1 \end{Bmatrix} = \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{Bmatrix}$$

where again we have simplified and rearranged the matrices using (a) through (d). All the terms in our expression for $\{A\}$ can now be

$$= \begin{Bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{Bmatrix} .$$

Using (b) we can rewrite this as

$$= 2\begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} .$$

These terms can now be evaluated using formulas (20) and (22) to yield

$$= 2[(1-2t)_+^{n-1} + (n-1)t(1-2t)_+^{n-2}] - (1-2t)_+^{n-1}$$

$$= (1-2t)_+^{n-1} + 2(n-1)t(1-2t)_+^{n-2} .$$

Example 5:

A modification of the previous example is to compute $\{A\}_t$ with

$$A = \begin{Bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{Bmatrix} .$$

Changing the dimension of A from 4 to 5 produces a considerable change in the complexity of the calculation and the form of the answer. When $n=5$, the probability $\{A\}_t$ has an interpretation similar to that in the previous example.

The distribution of the scan statistic (and related matters) has been extensively investigated by Naus and his co-workers Wallenstein, Huntington and Glaz.

Example 4:

We shall now compute

$$P\{S_1+S_2 > t, S_2+S_3 > t, S_3+S_4 > t, S_4+S_1 > t\} = \{A\}_t \quad \text{with}$$

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} .$$

When $n = 4$, (S_1, S_2, S_3, S_4) can be viewed as the spacings between four points placed uniformly on a circle with circumference equal to one and thus $\{A\}_t$ is the probability that no arc of length t contains more than two of the four points.

Let γ_i denote the $i^{\underline{th}}$ column of A. Since $\gamma_1 - \gamma_2 + \gamma_3 = \gamma_4$ and $1-1+1 = 1$, our theorem implies that

$$\{A\} = \{A_{1,\gamma_4}\} - \{A_{2,\gamma_4}\} + \{A_{3,\gamma_4}\}$$

$$= \begin{Bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{Bmatrix} - \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{Bmatrix} + \begin{Bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{Bmatrix} .$$

By applying property (a) and then property (d) to each term we obtain

Some Notation and Elementary Properties.

Let A be any matrix having at most n columns. Take Γ to be the matrix with n columns obtained by padding A with columns of zeros; $\Gamma = (A|0)$. Let k be the number of rows in A and define $\underset{\sim}{X} = (X_1, X_2, \ldots, X_k)'$ by $\underset{\sim}{X} = \Gamma \underset{\sim}{S}$. For any value of t we define

$$\{A\}_t = P\{X_i > t \text{ for all } i\} .$$

When the value of t is held fixed in an argument, we shall delete the subscript and just write $\{A\}$.

Our theorem and its consequences can sometimes be used to evaluate $\{A\}$. The following elementary properties are useful in the manipulations.

(a) If the $i^{\underline{th}}$ row of A *dominates* (is componentwise greater than or equal to) the $j^{\underline{th}}$ row, the $i^{\underline{th}}$ row can be deleted without changing the value of $\{A\}$.

The value of $\{A\}$ does <u>not</u> change when

(b) the columns are permuted,

(c) the rows are permuted, or

(d) a column of zeros is deleted.

The next three examples are calculations which in some simple cases (and with the appropriate choice of n) give the probabilities of certain types of clustering on a circle or interval. These sorts of calculations may be used to study the distribution of the scan statistic, a test for the presence of non-random clustering (see Naus (1966b) or Cressie (1977)).

The quantity $\binom{n-1}{r_1,r_2,\ldots,r_p}$ is the usual multinomial coefficient. The expression (22) is easily derived by direct methods which, however, do not apply in more complicated situations.

Example 3:

We now show how to compute $\mathbb{P}(\Gamma)$ when $\Gamma = (\gamma_{ij})$ is $2 \times n$ and $\gamma_{ij} = 0$ or 1 for all i and j. Let γ_i denote the $i^{\underline{th}}$ column of Γ. Define r, s and t to be the number of columns in Γ which are equal to $(1,0)'$, $(0,1)'$ and $(1,1)'$ respectively.

Since $\mathbb{P}(\Gamma)$ is invariant under permutation of the columns of Γ, $\mathbb{P}(\Gamma)$ depends only on r, s and t so that we may define $P(r,s,t) = \mathbb{P}(\Gamma)$. Assume $r > 0$, $s > 0$ and $t > 0$, then without loss of generality we may take $\gamma_1 = (1,0)'$, $\gamma_2 = (0,1)'$ and $\gamma_3 = (1,1)'$. Since $1+1-1 = 1$ and $\gamma_1+\gamma_2-\gamma_3 = (0,0)' = \underset{\sim}{0}$, our theorem gives

$$\mathbb{P}(\Gamma) = \mathbb{P}(\Gamma_{1,\underset{\sim}{0}}) + \mathbb{P}(\Gamma_{2,\underset{\sim}{0}}) - \mathbb{P}(\Gamma_{3,\underset{\sim}{0}}) .$$

This may be rewritten as

$$(23) \qquad P(r,s,t) = P(r-1,s,t) + P(r,s-1,t) - P(r,s,t-1) .$$

Repeated application of this recursion allows us to express $P(r,s,t)$ as a sum of distributions $P(i,j,k)$ in which $i = 0$ or $j = 0$ or $k = 0$. These distributions are easily handled by other methods. A generalization of (23) was obtained by a different method in Huffer (1982).

$$P\{Z_i > t_i \text{ for all } i\} = \xi[\underset{\sim}{x}_1|\underset{\sim}{x}_2|\cdots|\underset{\sim}{x}_p]$$

(21) where $$\xi(y_1|y_2|\cdots|y_p) = (Eh(y_1S_1|y_2S_2|\cdots|y_pS_p)) \prod_{i=1}^{p} y_i^{k_i-1}$$

$$= (1 - \sum_{i=1}^{p} t_i/y_i)_+^{n-1} \prod_{i=1}^{p} y_i^{k_i-1} .$$

When the components of the vectors $\underset{\sim}{x}_i$ are distinct, the divided differences can be evaluated by an obvious generalization of (14). When the components of each vector are all equal, we can calculate the divided differences by generalizing formula (15). For $1 \leq i \leq p$, let $\underset{\sim}{e}_i = (1,1,\ldots,1)$ have k_i components. It is easy to show that

$$\xi[y_1\underset{\sim}{e}_1|y_2\underset{\sim}{e}_2|\cdots|y_p\underset{\sim}{e}_p] = \left(\prod_{i=1}^{p} \frac{1}{(k_i-1)!} \left(\frac{\partial}{\partial y_i}\right)^{k_i-1}\right)\xi(y_1|y_2|\cdots|y_p) .$$

With ξ as given in (21), these partial derivatives are obtained by repeated use of (11). Setting $y_1 = y_2 = \cdots = y_p = 1$ in the resulting expression gives a formula for the joint distribution of nonoverlapping sums of uniform spacings.

(22)
$$P\{\sum_{j=m_{i-1}+1}^{m_i} S_j > t_i \text{ for } 1 \leq i \leq p\}$$
$$= \sum_{\substack{(r_1,r_2,\ldots,r_p) \\ 0 \leq r_i \leq k_i-1}} \binom{n-1}{r_1,r_2,\ldots,r_p}(1 - \sum_{i=1}^{p} t_i)_+^{n-1-\Sigma_i r_i} \prod_{j=1}^{p} t_j^{r_j}$$

where $m_0 = 0$ and $m_j = \Sigma_{i=1}^{j} k_i$ for $1 \leq j \leq p$. The summation is taken over all p-tuples of integers satisfying $0 \leq r_i \leq k_i-1$ for all i.

This result can easily be extended to handle $p > 2$. We shall state only the following special case. Assume Γ has a block diagonal form

$$\Gamma = \left(\begin{array}{cccc|c} \underset{\sim}{x}_1 & & & 0 & \\ & \underset{\sim}{x}_2 & & & 0 \\ & & \ddots & & \\ 0 & & & \underset{\sim}{x}_p & \end{array}\right)$$

where each $\underset{\sim}{x}_i$ is a row vector having k_i components. Let

$$\underset{\sim}{Z} = (Z_1, Z_2, \ldots, Z_p)' = \Gamma \underset{\sim}{S} ,$$

choose real numbers $t_1, t_2, \ldots, t_p$ and define

$$h(z_1 | z_2 | \cdots | z_p) = \begin{cases} 1 & \text{if } z_i > t_i \text{ for all } i , \\ 0 & \text{otherwise} , \end{cases}$$

so that $Eh(\underset{\sim}{Z}) = P\{Z_i > t_i \text{ for all } i\}$.

It is well known (see Feller volume II, problem 23 of chapter I) that

$$\text{(20)} \qquad P\{S_1 > t_1, S_2 > t_2, \ldots, S_p > t_p\} = (1 - \sum_{i=1}^{p} t_i)_+^{n-1} .$$

Repeated use of (16) followed by an application of (20) leads to the formula

$$\psi_k(x_1,x_2,\ldots,x_k) = \left(\frac{x_1}{x_1-x_2}\right)\psi_{k-1}(x_1,x_3,\ldots,x_k) + \left(\frac{x_2}{x_2-x_1}\right)\psi_{k-1}(x_2,x_3,\ldots,x_k) \ .$$

By induction as in example 1 we now obtain

$$\text{(16)} \qquad \begin{gathered} \psi_k(x_1,x_2,\ldots,x_k) = \phi[x_1,x_2,\ldots,x_k] \\ \text{with} \quad \phi(x) = x^{k-1}\psi_1(x) \ . \end{gathered}$$

Assume now that $p = 2$ and Γ has a more specialized form

$$\text{(17)} \qquad \Gamma = \left(\begin{array}{c|c|c} \underset{\sim}{x} & 0 & 0 \\ \hline 0 & \underset{\sim}{y} & 0 \end{array}\right)$$

where $\underset{\sim}{y}$ is an m-dimensional row vector. h is now a function of two arguments, $h = h(x|y)$. Applying (16) to the quantity $\psi_1(x)$ viewed as a function of $\underset{\sim}{y}$ yields

$$\text{(18)} \qquad \begin{gathered} \psi_1(x) = \theta[y_1,y_2,\ldots,y_m] \\ \text{with} \quad \theta(y) = y^{m-1}Eh(xS_1|yS_2) \ . \end{gathered}$$

Combining (16) and (18) leads to

$$\text{(19)} \qquad \begin{gathered} Eh(\Gamma\underset{\sim}{S}) = \xi[\underset{\sim}{x}|\underset{\sim}{y}] \\ \text{with} \quad \xi(x|y) = x^{k-1}y^{m-1}Eh(xS_1|yS_2) \end{gathered}$$

whenever Γ has the form in (17).

This is easily proved using formula (1) and the representation (13). When some of the values coincide, limiting arguments yield prescriptions (in terms of derivatives) for evaluating the divided differences. For example, when $x_i = x$ for all i and $y_j = y$ for all j, then

$$
\begin{aligned}
(15) \qquad & f[x_1,\ldots,x_p|y_1,\ldots,y_q] \\
& = \frac{1}{(p-1)!(q-1)!} \left(\frac{\partial}{\partial x}\right)^{p-1} \left(\frac{\partial}{\partial y}\right)^{q-1} f(x,y) \ .
\end{aligned}
$$

This result follows from (9) and (13).

<u>Example 2</u>:

Let Γ by any $p \times n$ matrix which can be written in the block form

$$
\Gamma = \left(\begin{array}{c|c|c} \underset{\sim}{x} & \underset{\sim}{b} & 0 \\ \hline 0 & \Lambda & 0 \end{array}\right)
$$

where $\underset{\sim}{x} = (x_1,x_2,\ldots,x_k)$ is a k-dimensional row vector and 0 denotes a matrix or vector of zeros. For an arbitrary bounded function h on $\mathbb{R}^p$ we define

$$
\psi_k(x_1,x_2,\ldots,x_k) = Eh(\Gamma \underset{\sim}{S}) \ .
$$

At first we shall consider $\underset{\sim}{b}$ and Λ to be fixed. The subscript k is included in the notation because we shall be varying the dimension of the vector $\underset{\sim}{x}$.

Assume that $x_1,x_2,\ldots,x_k$ are all distinct. Using our theorem exactly as in example 1 we obtain

Milne-Thomson, L.M. (1933). The Calculus of Finite Differences. Macmillan and Co., London.

Naus, J.I. (1966). Some Probabilities, Expectations, and Variances for the Size of Largest Clusters and Smallest Intervals. J. Amer. Statist. Assoc., 61, 1191 - 1199.

Naus, J.I. (1966b). A power comparison of two tests of non-random clustering. Technometrics, 8, 493 - 517.

Naus, J.I. (1982). Approximations for distributions of scan statistics. J. Amer. Statist. Assoc. 77, 177 - 183.

Pyke, R. (1965). Spacings (with discussion). J. Roy. Statist. Soc. Ser. B, 7, 395 - 449.

Varsi, G. (1973). The multidimensional content of the frustum of the simplex. Pacific J. Math. 46, 303 - 314.

Weisburg, H. (1971). The distribution of linear combinations of order statistics from the uniform distribution. Ann. Math. Statist. 42, 704 - 709.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 356	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Decomposing The Joint Distribution Of Several Linear Combinations Of Uniform Spacings	5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Fred Huffer	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0475
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Statistics, Stanford University, Stanford, CA 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Statistics & Probability Program Code 411SP.	12. REPORT DATE: April 25, 1985
	13. NUMBER OF PAGES: 27
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Spacings; uniform distribution; divided differences; clustering probabilities.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Using facts about divided differences, we prove an identity which can sometimes express the joint distribution of several linear combinations of uniform spacings as a sum of simpler distributions. This identity is useful in the exact computation of probabilities and expectations which arise in testing for uniformity. Examples are given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

7-85

DTIC